

# Take the sting out of data

---

## Addressing the 6 points of data pain

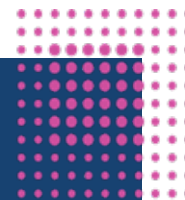


# Take the sting out of data:

## Addressing the 6 points of data pain

Digital data holds an incredible amount of value in today's world and without it, data science will suffer. Unfortunately, adding digital data to existing or new data science initiatives is often a massive struggle. Whether the data is an input for attribution, scoring models, machine learning, or otherwise, the challenges tend to remain the same:

- 1. Lack of a data model and schema**
- 2. Accuracy issues or missing key elements of data**
- 3. Inability to structure the digital data output in a usable format**
- 4. Inability to deliver timely data for models that require more immediate inputs**
- 5. Lack of individual level data**
- 6. Inability to build cross-session attributes tables at an individual level**



### Contents:

1. The 6 most common points of data pain
2. The future of data science
3. The Celebrus solution

# 1. Lack of a data model and schema

The importance of having a data model and schema readily available from whatever vendor you're using to capture digital data cannot be emphasized enough. It's a gap in the industry and something that enterprises typically don't realize is an issue until they try to connect that data from their marketing cloud or tech stack to an external system or vendor.

A comprehensive and [robust data model](#) provides the right data, in the right format, to drive real value from investments in data science and marketing technology. For effective data orchestration, the model must provide a consistent and logical framework for data and add value by intelligently pre-processing data into features and signals that are consumable by the wide array of systems that rely on that data to make decisions.

# 2. Accuracy issues or missing key elements of data

A key benefit of a customer data platform (CDP) is to build a unified customer view. Without accurate and complete data, it's not possible to realize this goal. Data science is dependent on extensive data to build reliable models, analytics, and actionable outputs. Traditional tag-based solutions are notoriously inflexible and error-prone, leading to bad data and missed opportunities.





A better alternative is to use a tagging-free solution that's deployed across sites and instantly captures every interaction without the need for extra data layers or tags. This enables complete data capture that can be used as and when needed, without ever having to worry about “we didn't tag for it” again.

### 3. Inability to structure digital data output in a usable format

Data captured without a relational structure is rendered relatively useless for data science without a heavy amount of costly transformation and joining of data downstream. Unfortunately, many CDPs are unable to structure data collected from various sources in a functional way.

One of the primary benefits of a CDP is to [reduce time spent managing and manipulating data](#). A modern customer data solution should capture data from multiple sources in a highly structured, yet lightweight format to properly feed downstream applications in a timely manner. Highly structured data also enables extraction of high-value data and slicing and dicing of data to support various business needs and use cases. Turning streams of clicks, views, and swipes into something that's instantly compatible with decisioning systems is a critical task that should be done at the data capture level to enable efficient and timely use of the data. For data scientists, this means more time spent on value-added activities such as analyzing data and creating predictive analytics, rather than spending time preparing data to feed models.



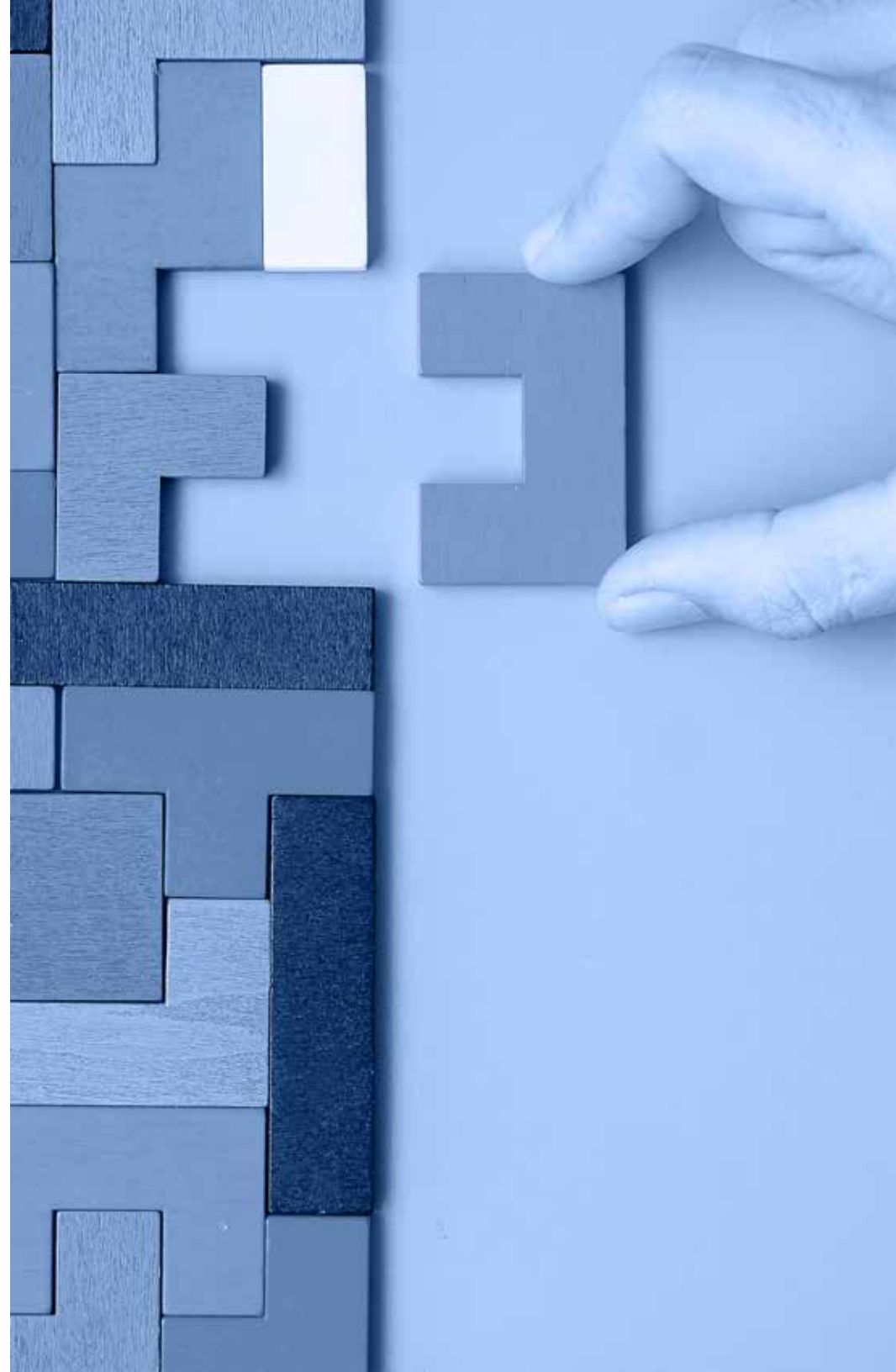
## 4. Inability to deliver timely data for models that require more immediate inputs

Many traditional data capture solutions don't provide [true real-time data](#). Activity information may be provided in seconds, but the addition of context typically delays the delivery of useful information by further seconds, minutes, or even hours. The data model is key to instantly detecting key signals of opportunity or threat from customer behavior and connecting this highly relevant dataset to a range of enterprise decisioning solutions. In short, it's about enabling faster reactions. If key signals aren't shared for seconds or worse, minutes or even hours, the opportunity may already be lost.

To change a webpage that someone is visiting in real-time, the capture, contextualization, and connection must occur within milliseconds.

## 5. Lack of individual level data

Personal identifiers, web behaviors, and transactions are essential data types to create detailed customer profiles at an individual level. Without individual data, use cases for personalization, segment hyper-targeting, and enhanced customer experience are null. Yet capturing identity and identifying channel visitors is often one of the biggest struggles for data scientists – especially for





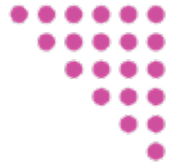
anonymous visitors. Many solutions lack the ability to identify users due to cookie challenges or browser restrictions. As a result, organizations are unable to persist identity in real-time.

The accurate understanding of identity is crucial to business decisions including personalization, marketing optimization, and fraud prevention, which means a [first-party identity graph](#) is vital.

## 6. Inability to build cross-session attribute tables at an individual level

Tag-based solutions, or those that rely on a data layer as a primary input, require advanced configuration to ensure that the micro-interactions within a particular page or experience are not only captured, but also tied to the session and to the individual. Even the most complex implementations struggle with the cross-session and device stitching of the individual. For data science, this presents as a core challenge given that many of these micro-interactions are required as signal inputs for the models being developed. Organizations often resort to using only a single channel or product line with a single decisioning system, losing out on the benefits of cross-channel customer understanding, and creating fragmented data silos that only add to the confusion.

To have a complete picture of customers, all data from all channels, for all products should be consolidated and available to any downstream application.



# The (better) future of data science

If you're a Data Scientist or analyst looking to leverage your valuable digital data in all data science initiatives, imagine the value to your organization and the time you could save by eliminating these typical challenges. The ideal data capture solution would look something like this:

- Zero tagging – instant, tag-free capture of all micro-interactions
- 80% reduction in the time you spend completing arduous data prep
- Easy access to a complete data model and schema processed in milliseconds for your use downstream, which you can expand upon as you see fit
- Instant stitching across sessions and devices
- Built-in compliance with all browser regulations, such as ITP, to ensure your attribution models and visitor identities persist beyond 7 days
- Ability to connect your data in a format of your choosing, using pre-built connectors for data science and machine learning out-of-the-box
- Automatic capture of individual digital profiles for all visitors to your digital properties
- Ability to structure outputs and have them available in milliseconds, in any format of your choosing based upon the specific models being built
- Rapid understanding of your digital behavioral data set to reduce time-to-value

**It's not an illusion or a far-fetched dream, this solution already exists.**

# Introducing Celebrus for data science

A data capture solution that combines 20+ years of passion with deep domain expertise in data capture. Our unique patented technology processes are quietly embedded in organizations who realized there's gold in digital customer data. Our global clients include large retail banks, airlines, insurance companies, and major retailers who are using Celebrus to feed their data ecosystem. Why? Because they crave truly 'real-time' data which enables them to deliver highly personalized, relevant, and 'in-the-moment' customer experiences. All this is achieved while simultaneously prioritizing customer privacy and meeting compliance regulations to safeguard their organization's reputation.

That's the true Celebrus advantage.

**CONNECT NOW**

Why settle for less? Reach out to start a conversation and learn how you too can evolve your approach to data science using our patented data solutions.

